



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/030,658A
Source: 1FW16 -
Date Processed by STIC: 3/4/04 -

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER:

10/030,658A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.

- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000

- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residu n or Xaa represents.

- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFW16

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/030,658A

DATE: 03/04/2004

TIME: 09:44:17

Input Set : A:\2004-02-25 4456-0101P.ST25.txt

Output Set: N:\CRF4\03042004\J030658A.raw

3 <110> APPLICANT: Yamamura Ken-ichi
4 Araki Kimi
6 <120> TITLE OF INVENTION: TRAP VECTORS AND GENE TRAPPING USING THE SAME
8 <130> FILE REFERENCE: 4456-0101P
10 <140> CURRENT APPLICATION NUMBER: 10/030,658A
11 <141> CURRENT FILING DATE: 2002-01-11
13 <150> PRIOR APPLICATION NUMBER: JP99/200997
14 <151> PRIOR FILING DATE: 1999-07-14
16 <160> NUMBER OF SEQ ID NOS: 17
18 <170> SOFTWARE: PatentIn Ver. 2.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 13
22 <212> TYPE: DNA
23 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
28 <400> SEQUENCE: 1
29 taccgttcgl ata 13
31 <210> SEQ ID NO: 2
32 <211> LENGTH: 13
33 <212> TYPE: DNA
34 <213> ORGANISM: Artificial Sequence
36 <220> FEATURE:
37 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
39 <400> SEQUENCE: 2
40 tatacgaacg gta 13
42 <210> SEQ ID NO: 3
43 <211> LENGTH: 34
44 <212> TYPE: DNA
45 <213> ORGANISM: Artificial Sequence
47 <220> FEATURE:
48 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
50 <400> SEQUENCE: 3
51 ataacttcgt atagcataca ttatacgaag ttat 34
53 <210> SEQ ID NO: 4
54 <211> LENGTH: 13
55 <212> TYPE: DNA
56 <213> ORGANISM: Artificial Sequence
58 <220> FEATURE:
59 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
61 <400> SEQUENCE: 4
62 ataacttcgt ata 13
64 <210> SEQ ID NO: 5

1. **John G. Thompson**
 2. **John H. Nieder**

RAW SEQUENCE LISTING

DATE: 03/04/2004

PATENT APPLICATION: US/10/030,658A

TIME: 09:44:17

Input Set : A:\2004-02-25 4456-0101P.ST25.txt

Output Set: N:\CRF4\03042004\J030658A.raw

65 <211> LENGTH: 13
 66 <212> TYPE: DNA
 67 <213> ORGANISM: Artificial Sequence
 69 <220> FEATURE:
 70 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
 72 <400> SEQUENCE: 5
 73 tatacgaagt tat 13
 75 <210> SEQ ID NO: 6
 76 <211> LENGTH: 34
 77 <212> TYPE: DNA
 78 <213> ORGANISM: Other *Invalid response - see item 10 on Error Summary Sheet*
 80 <220> FEATURE:
 81 <223> OTHER INFORMATION: Homologous recombination sequence
 83 <400> SEQUENCE: 6
 84 taccgttcgt atagcatata ttatacgaac ggta 34
 86 <210> SEQ ID NO: 7
 87 <211> LENGTH: 19
 88 <212> TYPE: DNA
 89 <213> ORGANISM: Artificial Sequence
 91 <220> FEATURE:
 92 <223> OTHER INFORMATION: Z1 forward primer used in PCR for B-geo detection
 94 <400> SEQUENCE: 7
 95 gcgttaccca acttaatcg 19
 97 <210> SEQ ID NO: 8
 98 <211> LENGTH: 18
 99 <212> TYPE: DNA
 100 <213> ORGANISM: Artificial Sequence
 102 <220> FEATURE:
 103 <223> OTHER INFORMATION: Z2 reverse primer used in PCR for B-geo detection
 105 <400> SEQUENCE: 8
 106 tgtgagcgag taacaacc 18
 108 <210> SEQ ID NO: 9
 109 <211> LENGTH: 22
 110 <212> TYPE: DNA
 111 <213> ORGANISM: Artificial Sequence
 113 <220> FEATURE:
 114 <223> OTHER INFORMATION: Ori2 forward primer used in PCR for detecting the
 replication
 115 origin region in pUC vector
 117 <400> SEQUENCE: 9
 118 gccagtggcg ataagtcgtg tc 22
 120 <210> SEQ ID NO: 10
 121 <211> LENGTH: 21
 122 <212> TYPE: DNA
 123 <213> ORGANISM: Artificial Sequence
 125 <220> FEATURE:
 126 <223> OTHER INFORMATION: Ori3 reverse primer used in PCR for detecting the
 replication
 127 origin region in pUC vector
 129 <400> SEQUENCE: 10
 130 cacagaatca ggggataacg c 21

RAW SEQUENCE LISTING

DATE: 03/04/2004

PATENT APPLICATION: US/10/030,658A

TIME: 09:44:17

Input Set : A:\2004-02-25 4456-0101P.ST25.txt

Output Set: N:\CRF4\03042004\J030658A.raw

132 <210> SEQ ID NO: 11
 133 <211> LENGTH: 400
 134 <212> TYPE: DNA
 135 <213> ORGANISM: Mus musculus
 137 <220> FEATURE:
 138 <221> NAME/KEY: misc_feature
 139 <222> LOCATION: (36)..(36)
 140 <223> OTHER INFORMATION: n is a, c, g, or t
 142 <220> FEATURE:
 143 <221> NAME/KEY: misc_feature
 144 <222> LOCATION: (70)..(70)
 145 <223> OTHER INFORMATION: n is a, c, g, or t
 147 <220> FEATURE:
 148 <221> NAME/KEY: misc_feature
 149 <222> LOCATION: (362)..(362)
 150 <223> OTHER INFORMATION: n is a, c, g, or t
 152 <220> FEATURE:
 153 <221> NAME/KEY: misc_feature
 154 <222> LOCATION: (364)..(364)
 155 <223> OTHER INFORMATION: n is a, c, g, or t
 157 <220> FEATURE:
 158 <221> NAME/KEY: misc_feature
 159 <222> LOCATION: (377)..(377)
 160 <223> OTHER INFORMATION: n is a, c, g, or t
 163 <400> SEQUENCE: 11

W--> 164 agaaacttaa acagcggata aacttcagtg atttanatca gagaagtatt ggaagtgtatt 60
 165 ctcaagggtan agcaacagcg gctaacaaca aacgtcagct tagtgaaaac cgaagccct 120
 166 tcaacttttt gcctalgcag attaatacta acaagagcaa ggatgctact gcaagtcttc 180
 167 caaagagaga gatgacaacg tcagcacagt gcaaagagtt gtttgcttct gctctaagta 240
 168 atgacctttt gcaaaactgt caatctctga agaagatggg agaggggagc ctgcatggga 300
 169 aacaccagat tgtaagcagg cttgttcaat cctgactata ttactaaagc tagttctatg 360
 170 cnanaagttt tgtaanaaaa atgaaagtct gcaatgttga 400
 172 <210> SEQ ID NO: 12
 173 <211> LENGTH: 416
 174 <212> TYPE: DNA
 175 <213> ORGANISM: Mus musculus
 177 <220> FEATURE:
 178 <221> NAME/KEY: misc_feature
 179 <222> LOCATION: (37)..(37)
 180 <223> OTHER INFORMATION: n is a, c, g, or t
 182 <220> FEATURE:
 183 <221> NAME/KEY: misc_feature
 184 <222> LOCATION: (363)..(363)
 185 <223> OTHER INFORMATION: n is a, c, g, or t
 187 <220> FEATURE:
 188 <221> NAME/KEY: misc_feature
 189 <222> LOCATION: (392)..(392)
 190 <223> OTHER INFORMATION: n is a, c, g, or t
 192 <220> FEATURE:

RAW SEQUENCE LISTING

DATE: 03/04/2004

PATENT APPLICATION: US/10/030,658A

TIME: 09:44:17

Input Set : A:\2004-02-25 4456-0101P.ST25.txt

Output Set: N:\CRF4\03042004\J030658A.raw

```

193 <221> NAME/KEY: misc_feature
194 <222> LOCATION: (401)..(401)
195 <223> OTHER INFORMATION: n is a, c, g, or t
197 <220> FEATURE:
198 <221> NAME/KEY: misc_feature
199 <222> LOCATION: (403)..(403)
200 <223> OTHER INFORMATION: n is a, c, g, or t
203 <400> SEQUENCE: 12
W--> 204 ttttctagct ttgcagcata aagcagagca agctatnagc tgtgatggat gactctgttg 60
205 ttacagaaac tacaggaagc ttatctggag tcagcatcac atctgaacta aatgaagaac 120
206 tgaatgattt aattcagcgl ttccataatc agcttcgtga ttctcagcct ccagctgttc 180
207 cagacaacag aagacaggca gaaagtcttt cattaactag agagatttct cagagcagaa 240
208 atccctcagt ttctgaacat ttacctgatg agaaagtaca qcttttttagc aaaatgagag 300
209 tactacagga aaagaacaag aaatggacaa attagttggg agaacttcat aaccttcgag 360
210 atnagcatct gaacaactca tcatttgtgc cntcaacttc ncnccaaaga agtggg 416
212 <210> SEQ ID NO: 13
213 <211> LENGTH: 484
214 <212> TYPE: DNA
215 <213> ORGANISM: Mus musculus
217 <220> FEATURE:
218 <221> NAME/KEY: misc_feature
219 <222> LOCATION: (33)..(33)
220 <223> OTHER INFORMATION: n is a, c, g, or t
222 <220> FEATURE:
223 <221> NAME/KEY: misc_feature
224 <222> LOCATION: (46)..(46)
225 <223> OTHER INFORMATION: n is a, c, g, or t
227 <220> FEATURE:
228 <221> NAME/KEY: misc_feature
229 <222> LOCATION: (54)..(54)
230 <223> OTHER INFORMATION: n is a, c, g, or t
232 <220> FEATURE:
233 <221> NAME/KEY: misc_feature
234 <222> LOCATION: (89)..(89)
235 <223> OTHER INFORMATION: n is a, c, g, or t
237 <220> FEATURE:
238 <221> NAME/KEY: misc_feature
239 <222> LOCATION: (244)..(244)
240 <223> OTHER INFORMATION: n is a, c, g, or t
242 <220> FEATURE:
243 <221> NAME/KEY: misc_feature
244 <222> LOCATION: (257)..(257)
245 <223> OTHER INFORMATION: n is a, c, g, or t
248 <400> SEQUENCE: 13
W--> 249 gtttctacac ctactgaaca gcagcagcca tttagctcaa aatccttnca gggnaaaaca 60
250 gagtatatgg cttttccaaa accctctgna aagcagttct tctcllggag cagaaaagca 120
251 aaggaatcaa gaaccagccc gaagaggaag ctgaaaacac taagacacca tggltatatg 180
252 atcaagaagg tggagtagaa aaaccatttt tcaagactgg atttacagag tctgtagaga 240
253 aagnlacaaa atagtanccg caaaaalcaa ccagatacaa gcaggagaaq acgtcgggtt 300

```

RAW SEQUENCE LISTING

DATE: 03/04/2004

PATENT APPLICATION: US/10/030,658A

TIME: 09:44:17

Input Set : A:\2004-02-25 4456-0101P.ST25.txt

Output Set: N:\CRF4\03042004\J030658A.raw

```

254 gatgaagaat cccttggaag gcttttagcag tatgcctgat cctatagacc caacatcagt 360
255 aactaaaaca tttaaaacaa gaaaagcacc tgcacaggcc agcctggcct ctaaggacaa 420
256 aactcccaaa tcaaagagta agaagaggat tctactcagc tgaagagtag agttaaaaat 480
257 attg 484
260 <210> SEQ ID NO: 14
261 <211> LENGTH: 211
262 <212> TYPE: DNA
263 <213> ORGANISM: Mus musculus
265 <400> SEQUENCE: 14
266 ctgtctgtca ttgtcgttct cctttagaag gcagaaaaga aatgggaaga aaaaaggcaa 60
267 aatctggaac actataacgg aaaggagttc gagaagctcc tggagggaagc tcaggccaac 120
268 atcatgaagt caattccaaa cctggagatg cccccagctt ccagcccagt glcaaaggga 180
269 gatgcggcag gggataaagt ggagctgtca g 211
272 <210> SEQ ID NO: 15
273 <211> LENGTH: 34
274 <212> TYPE: DNA
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
280 <400> SEQUENCE: 15
281 taccgttcgt atagcatata ttatacgaag ttat 34
283 <210> SEQ ID NO: 16
284 <211> LENGTH: 34
285 <212> TYPE: DNA
286 <213> ORGANISM: Artificial Sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
291 <400> SEQUENCE: 16
292 ataacttcgt atagcatata ttatacgaac ggta 34
294 <210> SEQ ID NO: 17
295 <211> LENGTH: 34
296 <212> TYPE: DNA
297 <213> ORGANISM: Artificial Sequence
299 <220> FEATURE:
300 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic DNA
302 <400> SEQUENCE: 17
303 tattgaagca talcgtatgt aatatgcttc aata 34

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 03/04/2004
PATENT APPLICATION: US/10/030,658A TIME: 09:44:18

Input Set : A:\2004-02-25 4456-0101P.ST25.txt
Output Set: N:\CRF4\03042004\J030658A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; N Pos. 36,70,362,364,377
Seq#:12; N Pos. 37,363,392,401,403
Seq#:13; N Pos. 33,48,54,89,244,257

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/030,658A

DATE: 03/04/2004

TIME: 09:44:18

Input Set : A:\2004-02-25 4456-0101P.ST25.txt
Output Set: N:\CRF4\03042004\J030658A.raw

L:164 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
M:341 Repeated in SeqNo=11
L:204 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
M:341 Repeated in SeqNo=12
L:249 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
M:341 Repeated in SeqNo=13